



Zhu, X., Zhang, Q., & Damian, M. F. (2016). Additivity of semantic and phonological effects: Evidence from speech production in Mandarin. *Quarterly Journal of Experimental Psychology*, 69(11), 2285-2304. <https://doi.org/10.1080/17470218.2015.1129427>

Peer reviewed version

License (if available):  
Other

Link to published version (if available):  
[10.1080/17470218.2015.1129427](https://doi.org/10.1080/17470218.2015.1129427)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is an Accepted Manuscript of an article published by Taylor & Francis in The Quarterly Journal of Experimental Psychology on 01/03/2016, available online:  
<http://www.tandfonline.com/10.1080/17470218.2015.1129427>

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

RUNNING HEAD: Additivity of semantics and phonology

**Additivity of semantic and phonological effects: Evidence from speech  
production in Mandarin**

Xuebing Zhu<sup>1</sup>, Qingfang Zhang<sup>2,3</sup>, and Markus F. Damian<sup>4</sup>

*<sup>1</sup>Institute of Linguistic Studies, Shanghai International Studies University, Shanghai,  
China*

*<sup>2</sup>Department of Psychology, Renmin University of China, China*

*<sup>3</sup>Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of  
Sciences, Beijing, China*

*<sup>4</sup>School of Experimental Psychology, University of Bristol, UK*

Word count: 7,821

Address correspondence to:

Qingfang ZHANG

Department of Psychology, Renmin University of China

59 Zhongguancun Street, Haidian District, Beijing 100872, China

Tel: 86-10-82501549

E-mail: qingfang.zhang@ruc.edu.cn

### Abstract

A number of previous studies using picture-word interference (PWI) tasks conducted with speakers of Western languages have demonstrated non-additive effects of semantic and form overlap between pictures and words, which may indicate underlying non-discrete processing stages in lexical retrieval. The present study used Mandarin speakers and presented Chinese characters as distractors. In two experiments, we crossed semantic relatedness with “pure” phonological (i.e., orthographically unrelated) relatedness, and found statistically additive effects. In a third experiment, semantic relatedness was crossed with orthographic overlap (phonological overlap was avoided), and once again we found an additive pattern. The results are discussed with regard to possible cross-linguistic differences between Western and non-Western languages in terms of phonological encoding, as well as concerning the locus of relatedness effects in PWI tasks.

*Key words:* speech production, semantic interference, phonological facilitation, lexical selection, Mandarin

Speaking involves the retrieval of semantic, syntactic, and phonological/phonetic information of the corresponding words. How these forms of information relate to each other remains a controversial issue in speech production theories (Bi, Xu, & Caramazza, 2009; Dell, 1986; Starreveld & La Heij, 1995, 1996a, 1996b). *Serial discrete* models (e.g., Levelt, Roelofs, & Meyer, 1999) argue that for a given target word, only a single selected lexical-semantic/syntactic node (“lemma”) spreads its activation to the phonological level, and semantic processing must be completed before phonological processing can begin. Non-serial models dispute some or all of these assumptions, and assume that phonological encoding can proceed based on partial and continuous input from the semantic level. *Cascaded* models (e.g., Humphreys, Riddoch, & Quinlan, 1988; Morsella & Miozzo, 2002) propose that multiple lexical-semantic candidates which are co-activated during retrieval of the target word transmit activation to the phonological level. *Interactive* models (e.g., Dell, 1986) additionally assume that transmission of activation between semantic and phonological encoding is bidirectional. In such models, phonological processing begins on the basis of early partial information provided by semantic processes, and partial phonological processing aids in retrieval at the earlier lexical-semantic level.

One way of empirically distinguishing discrete from non-discrete (cascaded or interactive) models is to apply “additive-factors logic” (Sternberg, 1969). Two experimental variables are manipulated and their effect on performance is measured. If the two variables show additive effects, then it can be concluded that they affect different and separate processing stages. By contrast, if the two variables show non-

additive effects, then either they act on a single processing stage, or they affect two processing stages but these two stages are closely related in terms of processing (i.e., they interact). Hence, according to additive factors logic, “factors that influence different stages will have additive effects on mean RT, whereas factors that influence stages in common will interact” (Sternberg, 1969, p. 282). This idea has been widely applied in experimental psychology, and also in the literature on word production.

One of the most widely used paradigms to investigate the cognitive processes involved in speech production is the picture-word interference task (hereafter PWI). Speakers name simple objects presented on a computer screen. Simultaneously or in close temporal proximity, “distractor” words are presented in printed or spoken form, and participants are instructed to ignore them as much as possible. Distractors still undergo processing, however, as evidenced by the general observation that a semantic relationship between a target picture and a distractor (e.g., picture: “sheep”, distractor: “goat”) slows response time relative to an unrelated condition (distractor: “chair”; e.g., Glaser & Döngelhoff, 1984). A further robust finding is that form overlap between picture and target (e.g., picture: “sheep”; distractor: “shoe”) speeds up responses relative to an unrelated condition (Schriefers, Meyer, & Levelt, 1990; Starreveld & La Heij, 1995). In PWI studies, it is common to manipulate the time interval between distractor and target onset (stimulus onset asynchrony, or SOA). By varying the point in time of distractor relative to target presentation, the distractor taps into successive stages of target preparation, hence potentially yielding information about relative temporal patterns.

Under the assumption that semantic effects in PWI reside at the lexical-semantic level whereas phonological effects are located at the phonological level, the employment of additive factors logic could indicate how the underlying processing stages relate to each other. To this aim, the usual semantically and phonologically related distractors are used, but additionally, so-called “mixed” distractors are included which are semantically as well as form-related (e.g., picture: “rabbit”; distractor: “rat”). In this way, semantic relatedness (semantically related, semantically unrelated) and phonological relatedness (phonologically relatedness, phonologically unrelated) could be factorially crossed. This allows to determine whether the two variables have additive or interactive effects according to “additive factors logic” (Sternberg, 1969). A number of studies have pursued this logic, and the general observation is non-additivity (Rayner & Springer, 1986; Starreveld & La Heij, 1995, 1996b; Taylor & Burke, 2002). This statistical interaction is found not only with printed (Starreveld & La Heij, 1995) but also with spoken distractors (Damian & Martin, 1999).<sup>1</sup> More specifically, the typically found pattern is that the semantic interference effect is attenuated when the distractors are also form-related to the target’s name; hence, “rabbit-rat” acts predominantly as a form-related pair whereas the semantic effect which should arise from shared category membership is much diminished (see Figure 1A for a set of results which is fairly representative of the broader pattern).

(Insert Figure 1 about here)

The observation of non-additivity between semantic and form-related distractor effects is seemingly at odds with serial theories of word production: if semantic and phonological stages are largely separated from each other (an assumption which constitutes the core of the seriality claim), it is not clear why semantic and form-related manipulations in PWI tasks would statistically interact. More specifically with regard to the semantically and form-related condition in PWI tasks, why does form overlap (presumably affecting the stage of phonological encoding) reduce the effects of semantic overlap (which presumably arises during lexical-semantic selection)? A possible explanation which has been advanced is that this pattern arises as a consequence of feedback from word-form encoding to lexical-semantic retrieval (e.g., Damian & Martin, 1999). Alternatively, both semantic and phonological effects might arise at the same processing level, namely word form retrieval (Starreveld & La Heij, 1995, 1996a, 1996b). Overall, the finding is generally taken to support non-discrete models of word production although this inference is not undisputed (e.g., Roelofs, Meyer & Levelt, 1996).

Recently, we (Zhu, Damian & Zhang, 2015) reported results from a study with a similar design but which showed a dramatically different pattern. Chinese Mandarin speakers named objects, and distractor words were either unrelated to the response, semantically related, form-related (they shared a spoken syllable with differing tone with the target response), or semantically and form-related. Hence, as in the previous studies summarised above, semantic and form relatedness were factorially crossed. Response latencies are shown in Figure 1B, and as is evident, both types of

relatedness showed almost perfect additivity. In addition to behavioural latencies, results from EEG (electroencephalography) were measured. Previous studies which had used EEG in conjunction with PWI studies had shown event-related potentials (ERPs) evoked by the two types of relatedness which appeared in roughly the same time window relative to picture onset (e.g., Dell'Acqua et al., 2010). By contrast, we found a serial pattern, with semantic effects appearing in a time window of 250-450 ms and phonological effects in a time window of 450-600 ms.

*Why are semantic and form effects in PWI tasks additive in Chinese, but not in Western languages?*

What could account for the discrepancy between the results obtained from speakers of Western languages, and Chinese? Zhu et al. (2015) discussed several possibilities. One possibility hinges on the origin of the form-related facilitation effects in PWI tasks. A common assumption is that form-related distractors contact their respective phonological representations. If these partially overlap with those of the target object, a portion of the target phonology (segments, and/or lexical entry) is pre-activated and easier to retrieve than in an unrelated condition (e.g., Roelofs, 1997). If so, such priming effects are diagnostic with regard to target preparation. However, it is possible that a facilitation effect could also arise from priming during distractor processing, rather than as a result of target preparation. A printed (or spoken) word might activate not only its own input code but also that of form-related neighbours (e.g., Marslen-Wilson & Warren, 1994; McClelland & Rumelhart, 1981). If so, a distractor which is form-related to the picture name might pre-activate the



target entry, and critically, this preactivation might not be related to form-encoding of the target, but it could reside at higher level (“lemma”) representation of the target.<sup>3</sup> In this case, it would be expected that semantic and form overlap statistically interact.

Whether “input priming” in PWI tasks contributes to facilitation effects is difficult to establish, and indeed some evidence argues against this possibility (e.g., Schriefers et al., 1990, Experiment 3). However, a further complication arises from the type of form relatedness. In languages with alphabetic scripts, phonology and orthography are generally confounded, and hence phonological overlap in a PWI task typically implies orthographic relatedness and vice versa. Does form-based facilitation in PWI arise as a consequence of orthographic, or phonological, overlap (or both)? A number of studies which have tried to disentangle orthographic from phonological influences suggest that both variables contribute independently (Posnansky & Rayner, 1978; Underwood & Briggs, 1984), and that orthographic facilitation might be larger than phonological facilitation (Lupker, 1982).

In languages with non-alphabetic orthographic systems such as Chinese, it is much easier to dissociate orthographic from phonological factors<sup>4</sup>. In accordance with the earlier studies with speakers of Western languages, both orthographic overlap (e.g., the target picture *bed*, 床, /chuang2/ paired with the orthographically similar, yet phonologically dissimilar distractor 庆 /qing4/, “celebration”) and phonological overlap (e.g., the same target paired with the distractor 创 /chuang4/, “creation”, which is orthographically unrelated but has the same spoken syllable) contribute to form-related facilitation. Bi et al. (2009) observed more substantial facilitation from

orthographic than from phonological overlap, and suggested an intriguing possibility: form priming in PWI might reflect a mixture of effects residing at multiple levels, with orthographically related distractors priming lexical (“lemma”) access, and phonologically related distractors affecting phonological encoding. Supporting evidence was reported by Zhang, Chen, Weekes, and Yang (2009) who factorially crossed phonological and orthographic relatedness and additionally manipulated SOA. Orthographic priming emerged at SOAs of -150 ms, 0 ms, and +150 ms, and phonological priming at SOAs of 0 ms and +150 ms. Furthermore, orthographic and phonological factors were additive at SOAs of -150 ms and 0 ms, but interacted at SOA = +150 ms. From these results, the authors inferred, in line with Bi et al.’s (2009) argument, that orthographic and phonological relatedness effects arise at independent processing levels: orthographically related distractors prime “lemmas” directly via the “input priming” account outlined above; phonologically related distractors affect phonological encoding proper. However, Zhao, La Heij, and Schiller (2012) in a similar study found larger facilitation for phonologically than for orthographically related distractors, and both types of relatedness appeared under comparable SOAs. From this the authors concluded that contrary to Bi et al. and Zhang et al.’s view, both orthographic and phonological overlap generate priming at the level of phonological encoding.

Overall, from extant findings, whether orthographic and phonological relatedness arise at different underlying processing stages remains controversial. Nevertheless, it remains a possibility that orthographic and phonological relatedness could result in

different types of priming in PWI tasks. Presumably, pure phonological overlap, as implemented in Zhu et al. (2015), might account for their central finding represented in Figure 1B: when orthographic overlap is avoided, form-related priming effect might strictly emerge from phonological encoding of the target, and in this case (and different from results with Western languages, such as in the results shown in Figure 1A), perfect statistical additivity is found.

An alternative possibility which could account for the discrepancy between previous findings from Western languages, and the novel results reported by Zhu et al. (2015), arises from the target language itself. It has recently been suggested (Chen, Dell, & Chen, 2004; O'Seaghdha, Chen, & Chen, 2010) that languages differ in the “proximate unit” of phonological encoding (i.e., the primary selectable unit below the word level, carrying particular salience as a speech planning unit). For speakers of Western languages, the proximate unit is assumed to be the phoneme. By contrast, in Chinese Mandarin, syllables have particular prominence as Mandarin has a limited inventory of syllables with relatively simple structure, and the orthographic system maps characters onto spoken syllables. Indeed, in computational work, Chen et al. (2004) demonstrated that syllables emerge from the statistics of Mandarin phonology, but not from English. Roelofs (2015) recently described detailed computational simulations of phonological encoding which contrasted Dutch/English with Mandarin (and Japanese). In the WEAVER framework (e.g., Levelt et al., 1999; Roelofs, 1992, 1997), the form network for Western languages such as English and Dutch encompasses three levels: 1] the morpheme corresponding to the target is activated, 2]

segments corresponding to the target name are chosen (labelled for order); simultaneously, a metrical frame is activated, 3] segments and metrical frame are merged into “syllable motor programs”. By contrast, four processing levels for Mandarin Chinese: 1] a morpheme corresponding to the target is activated, 2] atonal syllable nodes are activated; simultaneously, a tonal frame is activated, 3] segments are activated, 4] segments and tonal frames are merged into syllable motor programs. If this framework is accurate, phonological encoding for Mandarin speakers would involve a different processing structure than Western languages, and this might account for the additive relationship in Chinese.

#### *The current set of experiments*

The experiments reported below had a number of objectives. First, because the additive pattern reported by Zhu et al. (2015) represents an anomaly when compared to the extant findings obtained with speakers of Western languages, we wished to replicate and extend it. To this aim, Experiment 1 used a fairly similar design (i.e., semantic and “pure” phonological relatedness were factorially crossed) but added a range of SOAs, which were expected to provide potential information concerning the time course of the two types of relatedness. Experiment 2 attempted to amass further evidence via an experimental design in which the same distractor words, for each type of relatedness, were used across related and unrelated conditions. This addresses potential problems arising from the use of different distractor words across conditions, which is unavoidable in a design in which both types of relatedness are factorially crossed. Finally, in Experiment 3, we crossed semantic overlap with orthographic

(rather than phonological) relatedness. If, as suggested by Bi et al. (2009) and Zhang et al. (2009), orthographic and phonological overlap affect different levels of target processing, then we might predict, for Experiment 3, a statistical interaction between semantic and orthographic overlap. On the other hand, if a specific property of the target language (Chinese Mandarin) is responsible for the statistical additivity between semantic and form overlap, then the type of form overlap (Exp. 1 and 2: phonological; Exp. 3: orthographic) should be irrelevant and we should once again obtain additive effects of semantic and form overlap.

## Experiment 1

### Method

**Participants.** Thirty students at Beijing Forestry University and China Agricultural University (13 males, mean age 23) participated. All were native Mandarin Chinese speakers and reported normal or corrected-to-normal vision.

**Materials.** Eighty black and white line drawings were selected from Zhang and Yang's (2003) picture database: 16 as target pictures and 64 as fillers. Pictures were standardised to a size of approximately 6 x 6 cm and displayed at the center of the screen. Each target picture was paired with four distractor words. Distractor words were presented in 25-point Song font, centrally superimposed on the target pictures. A semantically related distractor word (condition S) was chosen that belonged to the same category as the target picture but had no phonological overlap (i.e., 眼睛, /yan3jing5/, "eye" as target, 嘴唇, /zui3chun2/, "lips" as distractor). A phonologically related word (P) was chosen that shared a syllable which always

differed in tone with the first character of the picture name (i.e., 沿海, /yan2hai3/, “coastal”). A semantically plus phonologically related word (SP) was chosen that belonged to same category as the target and shared a syllable with a different tone with the first character of the picture name (i.e., 咽喉, /yan1hou2/, “throat”). The P related and SP related distractor words had no orthographic relationship with target characters or the phonetic radicals of the first target characters in monosyllabic and disyllabic names. An unrelated word (U) was selected that stood in no obvious relationship to the target picture (i.e., 收据, /shou1ju4/, “receipt”). Among the 16 target pictures, six pictures had monosyllabic names, and ten pictures had disyllabic names, and distractor words always had corresponding numbers of syllables. Distractors in each condition were statistically matched for number of strokes and written frequency based on normative information reported in the database of Chinese Lexicon (2003). All items are reported in Appendix A. Sixty-four pictures were selected as fillers, and paired with unrelated distractor words. An additional twelve drawings were selected for the construction of practice and warm-up trials.

As to the relative degree of semantic relatedness in different conditions, a rating task as in Zhu et al. (2015) was carried out by sixteen native Mandarin speakers. Target picture name and distractor word pairs were rated on a 5-point scale, with 5 indicating that word pairs were highly semantically related and 1 indicating that word pairs were semantically unrelated. Rating scores in both S and SP conditions were high and close to each other, indicating the two related conditions were well-matched in terms of semantic overlap (Mean value: S =4.22; SP =4.12).

**Design.** The experimental design factorially crossed semantic relatedness (related vs. unrelated) and phonological relatedness (related vs. unrelated), and SOA (-150 ms, 0 ms, 150 ms) as within-participants and within-items variables. SOA was blocked, and the order of SOA blocks was counterbalanced according to a Latin square design. Within each block, a participant saw the 16 target pictures four times, for a total of 64 trials. Each SOA block additionally included 64 filler trials, resulting in 128 trials per block, and 384 trials for the entire experiment. The order of items within an SOA block was pseudo-randomised for each participant with the constraint that a particular target did not re-occur for at least five trials, and the first phoneme of a target name was never the same on consecutive trials.

**Apparatus.** The experiment was performed using E-Prime Professional Software (Version 1.1; Psychology Software Tools). Participants were seated approximately 60 cm from a computer screen. Naming latencies were measured from target onset using a voice-key, connected with the computer via a PST Serial Response Box.

**Procedure.** Participants were tested individually. They sat in a dimly lit room at a comfortable viewing distance in front of the computer. First, participants were asked to familiarise themselves with the experimental stimuli by viewing each target for 3000 ms with the correct name printed below. Then, participants were asked to name the pictures. Finally, the three experimental blocks were administered.

Each trial involved the following sequence: A fixation point (+) presented in the middle of the screen for 500 ms, followed by a blank screen for 500 ms. Then, the target picture appeared on the screen, and the distractor word was presented in a

manner appropriate for the particular SOA block (-150 ms: distractor onset preceded target onset by 150 ms; 0 ms: target and distractor appeared simultaneously; +150 ms: distractor onset followed target onset by 150 ms). Target pictures and distractor words disappeared when participants initiated a voice response. An inter-trial interval of 2,000 ms concluded each trial. The experiment took about 50 minutes in total.

## Results

Data from one participant were discarded due to a large number of hesitations and stutters. Latencies from incorrect responses (0.74%), other responses such as mouth clicks (0.41%), naming latencies longer than 2,000 ms or shorter than 200 ms (0.38%), and those deviating by more than three standard deviations (1.40%) were removed from all analyses. Table 1 presents the mean latencies and error percentages, presented by Distractor Type and SOA.

(Insert Table 1 about here)

The main objective of the experiment was to identify a potential interaction between semantic and phonological relatedness, under some or all SOAs. Figure 2 depicts the effects of semantic and phonological relatedness at each SOA separately. As can be seen, at SOA = -150 ms, only a semantic effect emerges. At SOA = 0 ms, both semantic and phonological effects are visible and show an almost perfectly additive relationship. The predicted effect based on the additivity assumption is 9 ms (semantic effect, 27 ms; phonological effect, -18 ms, the summation of those effects is 9 ms), and the observed effect is 2 ms. At SOA = +150 ms, neither variables exert much effect.



(Insert Figure 2 about here)

Analyses of variance (ANOVAs) were performed on response latencies, with SOA, semantic relatedness, and phonological relatedness as within-participants and within-items variables. The results are shown in Table 2. Both semantic and phonological relatedness interacted with SOA, suggesting that relatedness effects depended on the specific timing of target and distractor. Crucially, semantic and phonological relatedness did not statistically interact with each other,  $ps \geq .32$ , nor was the three-way interaction between semantic relatedness, phonological relatedness, and SOA significant,  $ps \geq .84$ . This implies that semantic and phonological variables exhibit an additive relationship.

(Insert Table 2 about here)

To further corroborate this inference, and despite the absence of a three-way interaction between semantic and phonological relatedness and SOA, we analysed the data separately for each SOA, with semantic and phonological relatedness as the variables. At SOA = -150 ms, the effect of semantic relatedness was significant,  $F(1, 28) = 11.26$ ,  $MSE = 891$ ,  $p < .01$ ,  $\eta_p^2 = .29$ ;  $F(1, 15) = 11.08$ ,  $MSE = 463$ ,  $p < .01$ ,  $\eta_p^2 = .43$ , but the effect of phonological relatedness was not,  $F_s \leq 0.94$ ,  $\eta_p^2 \leq .03$ , nor was the interaction between semantic and phonological relatedness,  $F(1, 28) = 0.04$ ,  $MSE = 481$ ,  $p = .84$ ,  $\eta_p^2 = .001$ ;  $F(1, 15) = 0.03$ ,  $MSE = 360$ ,  $p = .86$ ,  $\eta_p^2 = .002$ . At SOA = 0 ms, semantic relatedness was significant,  $F(1, 28) = 18.91$ ,  $MSE = 788$ ,  $p < .001$ ,  $\eta_p^2 = .40$ ;  $F(1, 15) = 16.22$ ,  $MSE = 559$ ,  $p < .01$ ,  $\eta_p^2 = .52$ , and so was phonological relatedness,  $F(1, 28) = 17.25$ ,  $MSE = 777$ ,  $p < .001$ ,  $\eta_p^2 = .38$ ;  $F(1, 15)$

$= 6.54$ ,  $MSE = 1125$ ,  $p < .05$ ,  $\eta_p^2 = .30$ , but the interaction was not,  $F1(1, 28) = 0.10$ ,  $MSE = 1202$ ,  $p = .76$ ,  $\eta_p^2 = .003$ ;  $F2(1, 15) = 0.31$ ,  $MSE = 509$ ,  $p = .59$ ,  $\eta_p^2 = .02$ .

Finally, at  $SOA = +150$  ms, neither semantic nor phonological relatedness were significant,  $F_s \leq 0.89$ ,  $p_s \geq .36$ ,  $\eta_p^2 \leq .05$ , nor was the interaction,  $F1(1, 28) = 1.86$ ,  $MSE = 386$ ,  $p = .18$ ,  $\eta_p^2 = .06$ ;  $F2(1, 15) = 0.64$ ,  $MSE = 386$ ,  $p = .44$ ,  $\eta_p^2 = .04$ .

The central finding concerns the absence of an interaction between semantic and phonological relatedness at the SOA under which both types of effects were present ( $SOA = 0$  ms). Bayesian analysis with the method suggested by Masson (2011) resulted in a Bayes factor of 5.20, with  $pBIC(H_0|D) = .839$  and  $pBIC(H_1|D) = .161$ , which according to the classification suggested by Raftery (1999) constitutes “positive” evidence for the null hypothesis (i.e., an additive pattern between semantic and phonological effects).

Error rates (overall less than 1%) were considered too low to allow for a meaningful statistical analysis.

## Discussion

In the current experiment, we found semantic interference at SOAs of -150 and 0 ms, but not at  $SOA = +150$  ms. This time course is fairly typical of numerous earlier studies in which semantic interference is also restricted to “earlier” SOAs but disappears under “later” SOAs (e.g., Schriefers et al., 1990). We found significant phonological priming at  $SOA = 0$  ms, but not at -150 ms or +150 ms. The observation of effects of phonological relatedness emerging strongly at  $SOA = 0$  ms is generally in line with previous PWI studies with visual distractors in Western languages (e.g.,

Damian & Martin, 1999; Starreveld & La Heij, 1996b) as well as with Chinese materials (Zhang et al., 2009; Zhao et al., 2012). In contrast to our study in which we found no phonological priming at SOA = -150 or +150, Zhao et al. reported weaker facilitation at SOA = -150 and +150 ms (12 and 8 ms facilitation respectively in Experiment 1, and 12 and 10 ms respectively in Experiment 2). In Zhang et al.'s (2009) study which was relatively similar to the current one, no phonological effect was found at SOA = -150, but 24 ms was found at SOA = +150 ms. Finally, in an experiment reported by Zhang and Weekes (2009), phonological effects emerged only under SOA = +100 ms, but not under SOA = 0 ms. The variability in effects across studies with regard to the exact time course of the effects doubtlessly reflects the complexity of the PWI paradigm, in which target and distractor dimensions each involve their own processing streams which need to be appropriately timed (via the SOA manipulation) to engage in cross-talk with one another. Most importantly, at the SOA under which both semantic and phonological effects were found (0 ms), the two types of relatedness showed a statistically additive pattern.

However, in the experimental design of Experiment 1, adopted from previous work such as Starreveld and La Heij (1995), it is necessary to use different distractor words across the experimental conditions. Although words were matched on written frequency and the number of strokes, other uncontrolled aspects of the distractor words could have affected processing times as well. For instance, Lupker (1979) proposed that the imageability of distractor words might affect naming latencies in a PWI task. Also, words varied in syntactic word class. It is conceivable that less

interference is obtained when target and context word do not belong to the same word class (see Mahon et al., 2007; Melinger & Koenig, 2007; Pechmann, Garrett, & Zerbst, 2004, for a recent debate).

For such reasons, a design would be preferable in which the same words can be used across related and unrelated conditions, and relatedness is manipulated by varying the pairing between targets and distractors. Experiment 2 aims to replicate the findings of Experiment 1 while improving on a number of aspects. Here, we used separate sets of distractor words for the semantic, phonological, and semantic and phonological condition; but critically, within each condition, a separate baseline was formed by recombining the distractors with the corresponding targets. This design allows for a better assessment of distractor effects of each type, i.e., within each condition separately (but not across conditions), the same distractor words are used.

## Experiment 2

### Method

**Participants.** Twenty-four students from the same participants pool as in Experiment 1 (10 males, mean age 23.0) took part. None of them had participated in Experiment 1.

**Materials** Fifteen black and white line drawings with disyllabic names were selected from Zhang and Yang's (2003) picture database. Each picture was paired with three types of distractor words: (1) a semantically related (S), but phonologically and orthographically dissimilar word; (2) a phonologically related (P), but orthographically dissimilar word; (3) a semantically plus phonologically related word

(SP). Each type of distractors was selected using the same criteria as in Experiment 1.

Distractors within each condition were then recombined with the picture names to form three unrelated conditions. All items are reported in Appendix B.

The same rating as in Experiment 1 was carried out to assess the degree of semantic relatedness. Across 15 participants, rating scores in both S and SP conditions were high and close to each other, indicating the two related conditions were well-matched in terms of semantic overlap ( $S = 4.45$ ;  $SP = 4.20$ ).

**Design.** The experimental design included type of relatedness (S, P, and SP related) and relatedness (related and unrelated) as within-participants and within-items variables. Each participant saw the 15 target pictures six times, resulting in a total of 90 trials. Pseudorandom sequences were generated by the same criteria as in Experiment 1.

**Apparatus and Procedure.** These were identical to Experiment 1.

## Results

Data from incorrect responses (0.70%), other responses such as mouth clicks (2.2%), naming latencies longer than 2,000 ms or shorter than 200 ms (0.74%), and those deviating by more than three standard deviations from the mean (1.85%) were removed from all analyses. Table 3 shows the mean latencies and error percentages, presented by distractor type.

(Insert Table 3 about here)

ANOVAs were conducted on the response latency means, with type of relatedness (S, P and SP) and relatedness (related and unrelated) as within-participants

and within-items variables. There was an effect of type of relatedness which reached significance in the analysis by items,  $F_2(2, 28) = 3.48$ ,  $MSE = 520$ ,  $p = .045$ ,  $\eta_p^2 = .20$ , but not in the analysis by subjects,  $F_1(2, 46) = 2.17$ ,  $MSE = 2419$ ,  $p = .126$ ,  $\eta_p^2 = .09$ . The effect of relatedness was not significant,  $F_1(1, 23) = .99$ ,  $MSE = 1197$ ,  $p = .330$ ,  $\eta_p^2 = .09$ ;  $F_2(1, 14) = .66$ ,  $MSE = 612$ ,  $p = .432$ ,  $\eta_p^2 = .05$ . The interaction between type of relatedness and relatedness was significant,  $F_1(2, 46) = 11.39$ ,  $MSE = 756$ ,  $p < .001$ ,  $\eta_p^2 = .33$ ;  $F_2(2, 28) = 6.69$ ,  $MSE = 860$ ,  $p = .004$ ,  $\eta_p^2 = .32$ .

We followed up the significant interaction between type of relatedness and relatedness with four separate contrasts. First, we tested for relatedness effects (i.e., unrelated vs. related), separately for each of the three types of relatedness. Results showed that the S related and unrelated conditions differed significantly from each other;  $F_1(1, 23) = 9.00$ ,  $MSE = 1385$ ,  $p = .006$ ;  $F_2(1, 14) = 5.15$ ,  $MSE = 1909$ ,  $p = .040$ , as did the P related and unrelated condition,  $F_1(1, 23) = 9.533$ ,  $MSE = 2323$ ,  $p = .005$ ;  $F_2(1, 14) = 13.79$ ,  $MSE = 923$ ,  $p = .002$ . By contrast, the SP related and unrelated conditions did not differ significantly,  $F_1(1, 23) = 1.30$ ,  $MSE = 1712$ ,  $p = .266$ ;  $F_2(1, 14) = .11$ ,  $MSE = 1832$ ,  $p = .422$ . Finally and most importantly, we assessed the degree of additivity between semantic and phonological relatedness with a formula (see Balota & Paul, 1996; Melinger & Abdel Rahman, 2004 for a similar logic) which on the left-hand side poses the SP effect (related minus unrelated), and on the right-hand side poses the sum of the S and P effects.

$$\text{SP effect} = \text{S effect} + \text{P effect}$$

If the effects of semantic and phonological effects are additive, then the two sides of the equation should be statistically equal; if the effects interact, then the two sides of the equation should deviate from zero. We tested this contrast via a coding of [-1, 1, -1, 1, 1, -1] across the S unrelated and related, P related and unrelated, and SP unrelated and related cells. This analysis did not return a significant result,  $F_1(1, 23) = .05$ ,  $MSE = 2118$ ,  $p = .830$ ;  $F_2(1, 14) = .11$ ,  $MSE = 4378$ ,  $p = .748$ , suggesting that semantic and phonological relatedness did not interact. Bayesian analysis resulted in a Bayes factor of 4.78, with  $pBIC(H_0|D) = .827$  and  $pBIC(H_1|D) = .173$ , which according to Raftery (1999) constitutes “positive” evidence for the null hypothesis of additive effects of semantic and phonological effects.

Error rates (overall less than 1%) were considered too low to allow for a meaningful statistical analysis.

## Discussion

In this experiment, we found both significant semantic interference and phonological facilitation, which replicates the classic effects found in Experiment 1 and in numerous previous studies. By contrast, a condition in which distractors were semantically and phonologically related showed very little effect, in line with the prediction that semantic interference and phonological facilitation cancelled each other out. Statistical tests indicated that semantic and phonological variables did not interact, but rather exerted additive effects. By forming separate baselines for each of the three related conditions in which the related distractor words were re-combined with target pictures to form unrelated pairings, we were able to exclude potential

confounds in terms of distractor properties (the variability in the “unrelated” latencies across the three types of distractors seen in Table 3 is probably attributable to such uncontrolled differences between the three distractor sets). In combination with the earlier findings by Zhu et al. (2015), Experiment 1 and 2 hence provide converging evidence for an additive pattern between semantic and phonological relatedness in Chinese spoken word production.

As outlined in the Introduction, the discrepancy between the findings from Western languages and Chinese might be caused by the fact that only in a non-alphabetic language such as Chinese it is possible to avoid orthographic overlap when manipulating phonological relatedness. Orthographic overlap, present in all previous experiments using Western languages, might have caused the statistical interaction between semantic and form overlap, but if orthographic overlap is avoided and distractors are exclusively related in terms of phonological overlap, then additivity might be found. In the third and final experiment, we explicitly tested this possibility by using a design very similar to Experiment 1, but now factorially crossing semantic with orthographic (rather than phonological) relatedness.

### **Experiment 3**

#### **Method**

**Participants.** Thirty students from the same participants pool as in Experiment 1 and 2 (5 males, mean age 21.6) took part. None of them had participated in Experiment 1 or 2.



**Materials.** Twenty-two black and white line drawings with disyllabic names were selected from Zhang and Yang's (2003) picture database. For each target picture, a semantically related distractor word (condition S) was chosen that belonged to the same category but had no orthographic or phonological overlap (i.e., 豹子, /bao4zi5/, "leopard" as a target, 大象, /da4xiang4/, "elephant" as a distractor). An orthographically related word (O) was chosen that shared a radical with picture name but semantically or phonologically dissimilar to the picture name (i.e., 约会, /yue1hui4/, "have a date" as an orthographically related distractor). A semantically plus orthographically related word (SO) was chosen that belonged to same category as the target and was shared a radical but phonologically dissimilar to the picture name (i.e., 豺狼, /chai2lang2/, "wolf"). Because there were two picture names with simple structure of Chinese character, the two of 22 O related and SO related distractors were chosen that shared visual similarity with target names. An unrelated word (U) was selected that stood in no obvious relationship to the target picture (i.e., 枕头, /zhen3tou2/, "pillow"). All items are reported in Appendix C.

The same semantic rating as in Experiment 1 was carried out. Across 11 participants, rating scores confirmed S and SO conditions were well-matched in terms of semantic overlap ( $S = 4.44$ ;  $SO = 4.34$ ). Similarly, as to orthographic overlap, an additional rating of orthographic overlap was carried out to access the degree of orthographic relatedness between distractors and targets. Rating scores for the O and SO conditions were high and close to each other, indicating the two related conditions were adequately matched in terms of orthographic overlap ( $O = 4$ ;  $SO = 3.6$ ).

**Design.** The experimental design factorially crossed semantic relatedness (related vs. unrelated) and orthographic relatedness (related vs. unrelated), and SOA (-150 ms, 0 ms, 150 ms) as within-participants and within-items variables. SOA was blocked, and the order of SOA blocks was counterbalanced according to a Latin square design. Within each block, a participant saw the 22 target pictures four times and 264 trials for the entire experiment. Pseudorandom sequences were generated by the same criteria as in Experiment 1.

**Apparatus and Procedure.** These were identical to Experiment 1 and 2.

## Results

Data from one participant were discarded due to a large number of hesitations and stutters. Latencies from incorrect responses (1.14%), other responses such as mouth clicks (2.16%), naming latencies longer than 2,000 ms or shorter than 200 ms (0.01%), and those deviating by more than three standard deviations (1.50%) were removed from all analyses. Table 4 presents the mean latencies and error percentages, presented by Distractor Type and SOA.

(Insert Table 4 about here)

Figure 3 depicts the effects of semantic and orthographic relatedness at each SOA separately. A semantic effect emerged only at SOA = 0 ms; an orthographic effect emerged at SOA = 0 ms and +150 ms. Hence, semantic and orthographic effects were simultaneously present only under SOA = 0 ms, and show an additive relationship. Under this SOA, the predicted effect based on the additivity assumption is -5 ms

(semantic effect, 24 ms; orthographic effect, -29 ms, the summation of those effects is -5 ms) and the observed effect is -1 ms.

(Insert Figure 3 about here)

Analyses of variance (ANOVAs) were performed on response latencies, with SOA, semantic relatedness, and orthographic relatedness as within-participants and within-items variables. The results are shown in Table 5. Both semantic and orthographic relatedness interacted with SOA, suggesting that relatedness effects depended on the specific timing of target and distractor. Crucially, semantic and orthographic relatedness did not statistically interact with each other,  $ps \geq .24$ , nor was the three-way interaction between semantic relatedness, phonological relatedness, and SOA significant,  $ps \geq .97$ . This implies that semantic and orthographic variables exhibit an additive relationship.

(Insert Table 5 about here)

As in Experiment 1, we further analysed the data separately for each SOA, with semantic and orthographic relatedness as the variables. At SOA = -150 ms, neither semantic nor orthographic relatedness were significant,  $F_s \leq 0.18$ ,  $ps \geq .67$ ,  $\eta_p^2 \leq .006$ , nor was the interaction,  $F1(1, 28) = .23$ ,  $MSE = 281$ ,  $p = .63$ ,  $\eta_p^2 = .008$ ;  $F2(1, 21) = 0.67$ ,  $MSE = 402$ ,  $p = .42$ ,  $\eta_p^2 = .03$ . At SOA = 0 ms, semantic relatedness was significant,  $F1(1, 28) = 21.67$ ,  $MSE = 868$ ,  $p < .001$ ,  $\eta_p^2 = .44$ ;  $F2(1, 21) = 10.68$ ,  $MSE = 1400$ ,  $p < .01$ ,  $\eta_p^2 = .34$  and so was orthographic relatedness,  $F1(1, 28) = 42.22$ ,  $MSE = 492$ ,  $p < .001$ ,  $\eta_p^2 = .60$ ;  $F2(1, 21) = 21.56$ ,  $MSE = 715$ ,  $p < .001$ ,  $\eta_p^2 = .51$ , but the interaction was not,  $F1(1, 28) = 0.34$ ,  $MSE = 314$ ,  $p = .57$ ,  $\eta_p^2 = .012$ ;

$F2(1, 21) = 0.32, MSE = 317, p = .58, \eta_p^2 = .015$ . At  $SOA = +150$  ms, the effect of orthographic relatedness was significant,  $F1(1, 28) = 14.42, MSE = 415, p < .001, \eta_p^2 = .34$ ;  $F2(1, 21) = 12.77, MSE = 309, p < .01, \eta_p^2 = .38$ , but the effect of semantic relatedness was not,  $F_s \leq 0.37, p_s \geq .55, \eta_p^2 \leq .013$ , nor was the interaction between semantic and orthographic relatedness,  $F1(1, 28) = 0.34, MSE = 456, p = .57, \eta_p^2 = .012$ ;  $F2(1, 21) = 1.06, MSE = 118, p = .31, \eta_p^2 = .048$ .

Bayesian analysis as carried out in Experiment 1 resulted in a Bayes factor of 4.53, with  $pBIC(H0|D) = .819$  and  $pBIC(H1|D) = .181$ , which according to Raftery (1999) constitutes “positive” evidence for the null hypothesis of additive effects of semantic and orthographic effects.

Error rates (overall 1.1%) were considered too low to allow for a meaningful statistical analysis.

## Discussion

We found semantic interference exclusively under  $SOA = 0$  ms (by contrast, in the first experiment, semantic effects were found at  $SOA = -150$  and  $0$  ms). This suggests that even with participants drawn from the same pool and very similar experimental setups and materials, there is some variability with regard to the time course of the effects. Orthographic facilitation effects were found at  $SOA = 0$  ms and  $+150$  ms. Earlier findings by Zhang et al. (2009) had shown orthographic priming which emerged at SOAs of  $-150$  ms,  $0$  ms, and  $+150$  ms. Zhao et al. (2012) showed the strongest orthographic priming at  $SOA = 0$  ms, and relatively less priming at more peripheral SOAs. Again we conclude that there is some unaccounted variability with

regard to the time course of semantic, orthographic, and phonological effects in PWI tasks. Nevertheless, in accordance with Zhu et al. (2015) and Experiment 1 and 2 above, both semantic and form overlap were simultaneously present at  $SOA = 0$  ms. And critically, in Experiment 3, semantic and orthographic overlap once again showed statistically additive effects. This suggests that the type of overlap (phonological in Exp. 1 and 2; orthographic in Exp. 3) is less relevant, and that with speakers of Mandarin and Chinese materials, semantic and form overlap exert an additive relationship in PWI experiments.

### **General Discussion**

In three PWI experiments carried out by native Mandarin speakers, we assessed the question of whether distractor words which are semantically or form-related to the picture name exert additive or non-additive effects. In Experiment 1, contrasting with previous studies conducted with speakers of Western languages but in line with results recently reported by Zhu et al. (2015), semantic and “pure” phonological relatedness simultaneously showed an effect ( $SOA = 0$  ms), they showed an additive relationship. Experiment 2 replicated the additive pattern with an improved relatedness manipulation in which the same distractor words could be used to form related and unrelated pairings, separately for each type of relatedness (semantic, phonological, semantic and phonological). Separate semantic and phonological effects were found, and again these formed a statistically additive pattern. In Experiment 3, we used a similar design to the first experiment but crossed semantic with orthographic (rather than phonological) overlap. Once again, under  $SOA = 0$  ms both

types of relatedness exerted simultaneous effects, and these were strictly additive. In combination, our findings highlight that when PWI experiments are conducted with speakers of Chinese Mandarin, semantic and form manipulations show additive effects, which contrasts with a number of similar studies with Western speakers in which the two variables always interacted (e.g., Damian & Martin, 1999; Rayner & Springer, 1986; Starreveld & La Heij, 1995, 1996b; Taylor & Burke, 2002).

Given that our central finding constitutes a null result (an absence of an interaction, although in the context of highly significant main effects) it might be argued that our studies suffered from insufficient statistical power. However, Bayesian analysis suggested a reasonable amount of support for the null interaction. Taken together, we are confident that in our studies, the empirically obtained additivity was genuine, rather than due to a lack of power to detect an interaction.

What are the implications of these results? A possibility first raised by Bi et al. (2009) and also endorsed in Zhang et al. (2009) is that orthographic and phonological relatedness effects in PWI arise at different loci, with the former largely affecting the stage of lexical-semantic (“lemma”) retrieval, and only the former rendering effects which genuinely reside at the level of target phonological encoding. If so, one might argue that semantic and form relatedness interact in PWI studies with Western speakers because here, there is an unavoidable confound between sound and spelling. By contrast, with Chinese materials, this confound can be avoided, as was the case in Zhu et al. (2015) and our Experiments 1 and 2. However, Experiment 3 manipulated

form overlap in terms of orthography and found a similar additive pattern. We conclude that the type of overlap is unlikely to be the critical variable.

A more likely candidate is the target language itself. Figure 4 shows a scheme, adapted from Zhu et al. (2015) and based on Roelofs' (2015) adaptation of the WEAVER framework to non-alphabetic languages. In our Experiment 1 and 2, phonologically related distractors shared a syllable with the target name, but the shared syllable always differed in tone. Hence, the most plausible locus of phonologically based facilitation effects in PWI tasks is the processing level of the "atonal syllable". Further supporting this inference is the finding that with Chinese speakers, segmental overlap alone generates little or no priming (Wong & Chen, 2008, 2009) thus the segmental level is unlikely to be the locus of facilitation. Hence, for distractor words in the phonologically related condition, the most likely scenario is that they are rapidly recoded into phonological format (which is in line with the claim that reading of Chinese characters critically involves phonology; e.g., Perfetti & Tan, 1998; Spinks, Liu, Perfetti & Tan, 2000) and that the shared atonal syllable is pre-activated and hence easier to encode when the target response is planned. In our Experiment 3, the orthographically related distractors shared a radical with the target name. To account for priming effects from these distractors, one could assume that a cohort of orthographically related items is co-activated in the orthographic lexicon, and that these activate their corresponding entries in the phonological lexicon (indeed, these assumptions are very similar to the model outlined by Zhao et al., 2012). Again, form-related priming would take place at the level of the (atonal) syllable. Hence,

both phonologically and orthographically based facilitation effects would have a common locus: a syllabic representational layer. This scenario would of course not answer the question of why the purported locus of the effects is “encapsulated” from earlier lexical-semantic retrieval in a way that it evidently isn’t in Western languages. An account in terms of “proximate units” (i.e., syllables are of crucial prominence for speakers of Chinese, and because of the relatively small numbers of relevant spoken syllables, these might be pre-stored) appears tentative and to some extent post hoc.

Yet another possibility is that phonologically and orthographically related distractors exert their effects at the same, shared locus, but that this is at the level of “syllabic motor programs”, rather than at the atonal syllabic level (see Figure 4). It is currently believed (e.g., Cholin, Levelt & Schiller, 2006; Levelt & Wheeldon, 1994) that phonological encoding proper (i.e., access to lexical phonological entries and corresponding segments; stress pattern etc.) is followed by access to a “mental syllabary” in which articulatory “gestural scores” for frequent syllables are pre-specified. For spoken Mandarin, such a syllabary should be particularly prominent: Chinese has relatively few syllable types of comparatively simple phonological complexity, clear syllable boundaries, and no resyllabification. If the syllabary is the locus of form-based priming effects in PWI tasks, it would not be implausible that processing is discrete with regard to the preceding stage of phonological encoding, and this would account for the additive pattern found in our experiments. Such a scenario would also account for the fact that EEG results reported in Zhu et al. (2015) showed a time window for the phonological effect which is quite “late” relative to



picture onset (450-600 ms). However, it is *prima facie* not clear why phonologically related distractors in our Experiments 1 and 2 generated priming, given that they shared a syllable but differed in tone; presumably, at the level of the syllabary, syllabic representations should be tonally specified. Perhaps entries at the syllabary level can be partially primed even by matching syllables which differ in tone.

In sum, we investigated the relationship between semantic and form-related variables with three PWI experiments by using Chinese characters as distractors. Semantic and phonological relatedness exerted additive effects in Experiments 1 and 2, and semantic and orthographic relatedness similarly showed an additive pattern in Experiment 3. This additive pattern contrasts with a number of previous findings obtained with speakers of Western languages. Interpreting the additivity between semantic and form-based relatedness effects depends on a number of assumptions regarding how and where such effects arise. At minimum, however, our results provide crucial constraints for current and future computational theories of word production. The discrepancy between PWI studies conducted with speakers of Western languages (in which semantic and form-based priming interact) and of Chinese (in which the two forms of relatedness exert additive effects) is significant and will ultimately need to be accounted for within a unified theoretical framework of word production.

### **Acknowledgments**

This research was supported by the National Natural Science Foundation of China under Grant (31170977, 31471074), the Fundamental Research funds for the Central Universities, and the Research Funds of Renmin University of China under Grant (No. 14NXLF12) to Qingfang Zhang.

## References

- Balota, D. A., & Paul, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 827–845.
- Bi, Y. C., Xu, Y. D., & Caramazza, A. (2009). Orthographic and phonological effects in the picture-word interference paradigm: Evidence from a logographic language. *Applied Psycholinguistics*, 30, 637–658.
- Chen, T.-M., Dell, G. S., & Chen, J.-Y. (2004). A cross-linguistic study of phonological units: Syllables emerge from the statistics of Mandarin Chinese but not from the statistics of English. *Proceedings of the 26<sup>th</sup> Annual conference of the Cognitive Science Society*, 216–220.
- Chinese Linguistic Data Consortium. (2003). 现代汉语通用词表 [Chinese lexicon] (CLDC-LAC-2003-001). Beijing, China: Tsinghua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation.
- Cholin, J., Levelt, W. J., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, 99, 205–235.
- Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 345–361.
- DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii Press.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production.

*Psychological Review*, 93, 283–321.

Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., & Grainger, J.

(2010). ERP evidence for ultra-fast semantic processing in the picture–word interference paradigm. *Frontiers in psychology*, 1:177.

Glaser, W. R., & Döngelhoff, F. J. (1984). The time course of picture-word

interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), 640–654.

Humphreys, G.W., Riddoch, M.J., & Quinlan, P.T. (1988). Cascade processing in

picture identification. *Cognitive Neuropsychology*, 5, 67–103.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental

syllabary? *Cognition*, 50, 239–269.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in

speech production. *Behavioral and Brain Sciences*, 22, 1–75.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word

interference task. *Memory & Cognition*, 7, 485–495.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture-word

interference. *Canadian Journal of Psychology*, 36, 349–367.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007).

Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm.

*Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 503–535.

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, 101, 653–675.

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, 43, 679–690.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375–407.

Melinger, A., & Abdel Rahman, R. (2004). Investigating the interplay between semantic and phonological distractor effects in picture naming. *Brain and Language*, 90, 213–220.

Melinger, A., & Koenig, J.-P. (2007). Part-of-speech persistence: The influence of part-of-speech information on lexical processes. *Journal of Memory and Language*, 56, 472–489.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 1146–1160.

- Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 555–563.
- O'Seaghdha, P. G., Chen, J.-Y., & Chen, T.-M. (2010). Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition*, 115, 282–302.
- Pechmann, T., & Garrett, M., & Zerbst, D. (2004). The time course of recovery for grammatical category information during lexical processing for syntactic construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 723–728.
- Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 101–118.
- Posnansky, C. J., & Rayner, K. (1978). Visual vs. phonemic contributions to importance of initial letter in word identification. *Bulletin of the Psychonomic Society*, 11, 188–190.
- Raftery, A. E. (1999). Bayes factors and BIC: Comment on "A critique of the Bayesian information criterion for model selection". *Sociological Methods & Research*, 27, 411–427.
- Rayner, K., & Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture word interference task. *British Journal of Psychology*, 77, 207–222.

- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284.
- Roelofs, A. (2004). Comprehension-based versus production-internal feedback in planning spoken words: A rejoinder to Rapp and Goldrick (2004). *Psychological Review*, 111, 579–580.
- Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research*, 57, 22–37.
- Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 246–251.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: The role of phonological information. *Cognition*, 76, B1–B11.
- Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 686–698.

- Starreveld, P. A., & La Heij, W. (1996a). The locus of orthographic-phonological facilitation: Reply to Roelofs, Meyer, and Levelt (1996). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 252–255.
- Starreveld, P. A., & La Heij, W. (1996b). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 896–918.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276–315.
- Taylor, J., & Burke, D. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psychology and Aging*, 17, 662–676.
- Underwood, G., & Briggs, P. (1984). The development of word recognition processes. *British Journal of Psychology*, 75, 243–255.
- Wong, A. W. K., & Chen, H.-C. (2008). Processing segmental and prosodic information in Cantonese word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1172–1190.
- Wong, A. W. K., & Chen, H. -C. (2009). What are effective phonological units in Cantonese spoken word planning? *Psychonomic Bulletin & Review*, 16, 888–892.
- Zhang, Q, Chen, H. -C., Weekes, B. S., & Yang, Y. (2009). Independent effects of orthographic and phonological facilitation on spoken word production in Mandarin. *Language Speech*, 52, 113–126.



- Zhang, Q., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes*, 24, 1082–1096.
- Zhang, Q., & Yang, Y. (2003). The determiners of picture naming latency. *Acta Psychologica Sinica*, 35, 447–454 (in Chinese).
- Zhao, H., La Heij, W., & Schiller, N. O. (2012). Orthographic and phonological facilitation in speech production: New evidence from picture naming in Chinese. *Acta Psychologica*, 139, 272–280.
- Zhou, Y. G. (1978). To what degree are the “phonetics” of present day Chinese characters still phonetic? *Zhongguo Yuwen*, 3, 172–177.
- Zhu, X., Damian, M. F., & Zhang, Q. (2015). Seriality of semantic and phonological processes during overt speech in Mandarin as revealed by event-related brain potentials. *Brain and Language*, 144, 16–25.

### Footnotes

<sup>1</sup>The only exception to this universal pattern that we are aware of was reported in a study by Melinger and Abdel Rahman (2004) in which, rather than using “mixed” distractors, two distractor words were simultaneously presented with the target picture, and hence in the mixed condition, semantic and form information was delivered by two separate elements (e.g., the target “pig” paired with the two distractors “pill” and “dog”). In this variant of the task, semantic and form-based effects were statistically additive. It was proposed that PWI effects are sensitive to how the information is introduced into the system. Specifically, in the single mixed distractor case, semantic and phonological codes engage (via feedback) in a form of mutual “resonance” such that lexical entry and phonological form activate each other. By contrast, with double distractors, distractor representations cannot benefit from this form of co-activation since semantically-related and form-related codes correspond to two separate entries.

<sup>2</sup>However, it should be noted that Meyer and Schriefers (1991) showed form-related priming effects in PWI tasks from distractors that overlapped in the word-final position, which suggests that cohorts are probably not the only cause of form-related facilitation.

<sup>3</sup>This line of reasoning was taken by Roelofs, Meyer, and Levelt (1996) in order to account for the non-additive effects reported by Starreveld and La Heij (1995; 1996b): the mixed distractor calf activates the lemma cat by means of a direct route,

or by means of phonological input representation (indirect route). Hence, a mixed distractor will activate its own lemma and also that of the target lemma cat. The same holds for an orthographically related distractor (i.e., cap) to the target cat (Roelofs, 2004; Roelofs, Meyer, & Levelt, 1996). In this case, as argued by Roelofs et al. (1996), the finding of non-additivity between semantic and form relatedness in PWI tells us very little about how the processing stages of lexical-semantic retrieval and phonological encoding relate to each other.

<sup>4</sup>Concerning the claim that the Chinese orthographic system dissociates sound from spelling, two aspects deserve highlighting. First, Chinese characters consist of sub-components (“radicals”) which provide probabilistic information about meaning and pronunciation of the underlying character (Zhou, 1978). DeFrancis (1989) emphasized that Chinese orthography is a speech-based script since more than 85% of Chinese characters are phonograms. This could be seen as at odds with our claim that by using Chinese materials, we are able to fully dissociate sound from spelling. However, in the system of Chinese characters, numerous homophones exist for a syllabic pinyin (regardless of tone), and these homophones are typically orthographically unrelated. For example, for the syllabic pinyin /zha/, corresponding Chinese characters could be 虎 (/hu3/), 狐 (/hu2/), or 胡 (/hu2/), with no orthographic overlap between them. Therefore, for our phonologically related or semantically plus phonologically related distractors, orthographic overlap with the target names could be fully avoided. Second, younger Chinese individuals will have

also been taught pinyin, a romanised transcription system of Chinese characters, so will have had considerable exposure to alphabetic scripts.

**Table 1**

Mean picture naming latencies (in ms) and error percentages (in parentheses) by

Distractor Type and Stimulus-Onset Asynchrony (SOA) in Experiment 1

Distractor Type	SOA (in ms)		
	-150	0	+150
Semantic	747 (1.29)	770 (1.29)	705 (0.43)
Phonological	732 (0.86)	725 (0.43)	704 (0.22)
Semantic & Phonological	749 (1.08)	745 (0.86)	703 (0.86)
Unrelated	728 (0.43)	743 (0.65)	696 (0.43)
Semantic effect	19*	27*	9
Phonological effect	4	-18*	8
Sem. &Phon. effect	21**	2	7

\* $p < .05$ . \*\* $p < .01$ .

**Table 2**

Analysis of Variance with Stimulus-Onset Asynchrony (SOA), Semantic Relatedness (S) and Phonological Relatedness (P) of Experiment 1.

Source	Participants					Items			
	df1	df2	MSE	F1	$\eta_p^2$	df2	MSE	F2	$\eta_p^2$
SOA	2	56	5373	12.00***	0.3	30	1026	34.05***	0.69
S	1	28	759	26.93***	0.49	15	427	26.39***	0.64
P	1	28	750	3.20†	0.1	15	641	2.21	0.13
SOA x S	2	56	821	3.14*	0.1	30	459	3.51*	0.19
SOA x P	2	56	507	11.36***	0.29	30	603	5.11*	0.25
S x P	1	28	586	1.01	0.04	15	488	0.68	0.04
SOA x S x P	2	56	741	0.18	0.006	30	383	0.11	0.007

† $p < .10$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

**Table 3**

Mean picture naming latencies (in ms), standard errors (SE) and errors percentage (in parentheses) by Distractor type in Experiment 2.

Relatedness	Distractor Type								
	S			P			SP		
	RT	SE	Err	RT	SE	Err	RT	SE	Err
Related	704	15	0.56	665	15	0.0	689	15	1.67
Unrelated	681	15	0.28	696	13	1.11	699	16	0.83
Effect	23**			-31**			-10		

NOTE: \*\* $p < .01$ . S = semantically related, P = phonologically related, SP = semantically and phonologically related.

**Table 4**

Mean picture naming latencies (in ms) and error percentages (in parentheses) by

Distractor Type and Stimulus-Onset Asynchrony (SOA) in Experiment 3

Distractor Type	SOA (in ms)		
	-150	0	+150
Semantic	662 (1.72)	692 (1.41)	634 (1.57)
Orthographic	660 (0)	639 (0.63)	621 (1.25)
Semantic & Orthographic	662 (1.25)	667 (2.51)	621 (0.94)
Unrelated	663 (0.31)	668 (1.41)	638 (0.63)
Semantic effect	-1	24**	-4
Orthographic effect	-3	-29**	-17**
Sem. & Ortho. effect	-1	-1	-17**

\* $p < .05$ . \*\* $p < .01$ .



**Table 5**

Analysis of Variance with Stimulus-Onset Asynchrony (SOA), Semantic Relatedness (S) and Orthographic Relatedness (P) of Experiment 3.

Source	Participants					Items			
	df1	df2	MSE	F1	$\eta_p^2$	df2	MSE	F2	$\eta_p^2$
SOA	2	56	2565	19.16***	0.41	42	434	87.40***	0.81
S	1	28	644	8.24***	0.23	21	1055	4.09*	0.16
O	1	28	338	51.84***	0.65	21	769	15.51***	0.43
SOA x S	2	56	413	16.52***	0.37	42	726	7.41**	0.26
SOA x O	2	56	455	10.20***	0.27	21	315	11.82***	0.36
S x O	1	28	450	0.70	0.02	21	324	1.47	0.07
SOA x S x O	2	56	300	0.02	0.001	42	257	0.04	0.002

† $p < .10$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

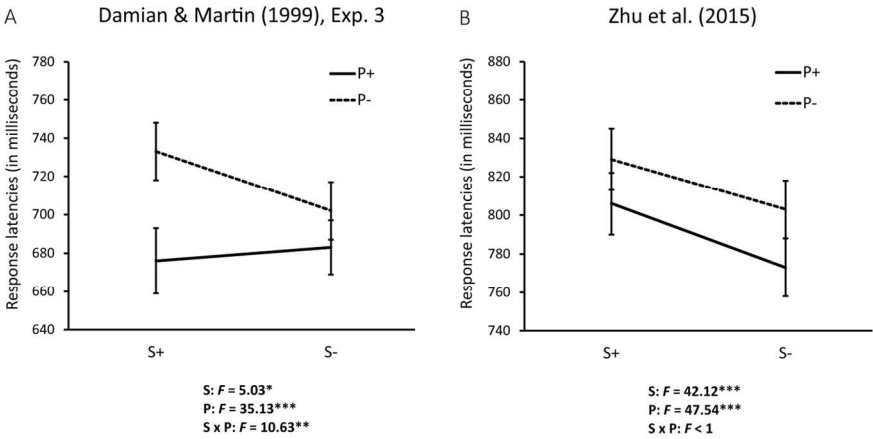
### Figure Captions

*Figure 1.* Results from Damian and Martin (1999; Experiment 3, SOA = 0 ms) and Zhu, Damian and Zhang (2015). S+ and S- indicate semantically related and unrelated conditions, respectively; P+ and P- indicate phonologically related and unrelated conditions, respectively.  $*p < .05$ ;  $**p < .01$ ;  $***p < .001$ .

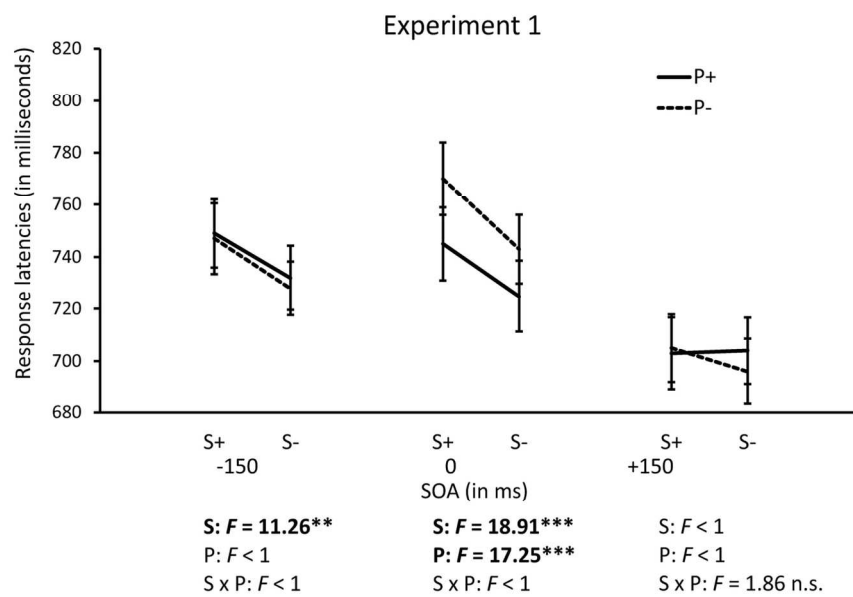
*Figure 2.* Experiment 1. Semantic and phonological effects and their relationship as a function of stimulus-onset asynchrony (SOA). S+ and S- indicate semantically related and unrelated conditions, respectively; P+ and P- indicate phonologically related and unrelated conditions, respectively.  $*p < .05$ ;  $**p < .01$ ;  $***p < .001$ .

*Figure 3.* Experiment 3. Semantic and orthographic effects and their relationship as a function of stimulus-onset asynchrony (SOA). S+ and S- indicate semantically related and unrelated conditions, respectively; O+ and O- indicate orthographically related and unrelated conditions, respectively.  $*p < .05$ ;  $**p < .01$ ;  $***p < .001$ .

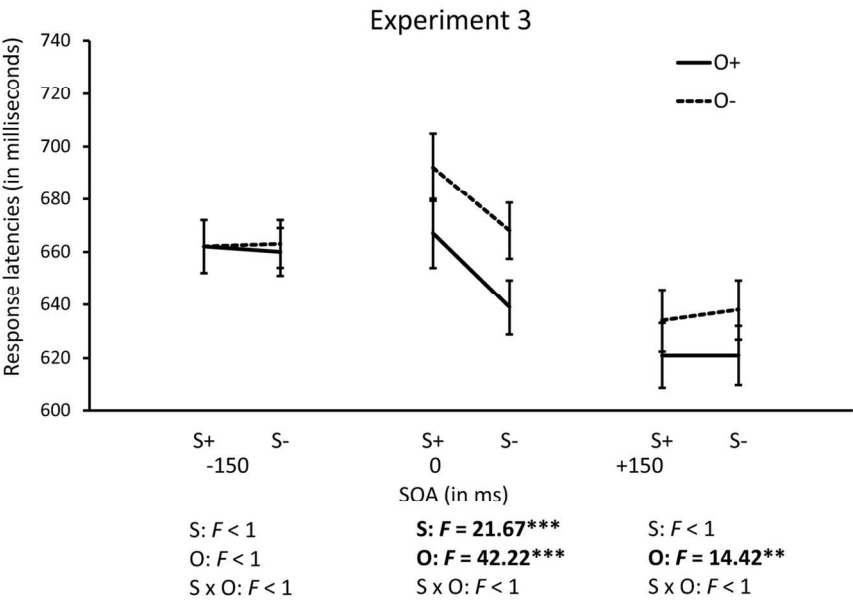
*Figure 4.* Hypothetical processing stages involved in phonological encoding of the concept “tiger” in Mandarin.



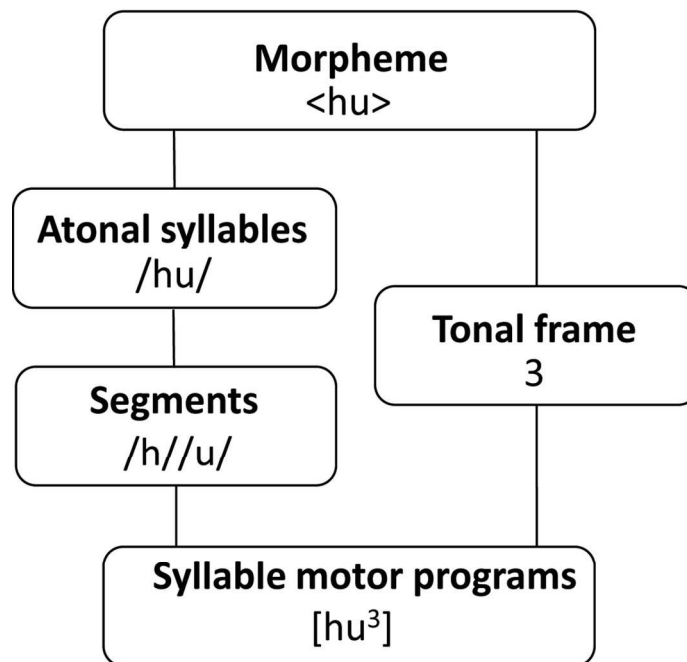
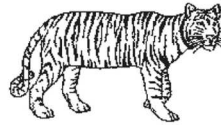
139x68mm (300 x 300 DPI)



131x79mm (300 x 300 DPI)



130x80mm (300 x 300 DPI)



144x136mm (300 x 300 DPI)